DOCUMENT RESUME

ED 103 448 TM 004 279

AUTHOR Remer, Rory
TITLE The Development of an Instrument to Provide Performance Predictive Information for Counselor Trainee Selection: Phase I.
PUB DATE Jan 73
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Counselor Qualifications; *Counselor Training; Criteria; Evaluation; Interpersonal Relationship; Measurement Techniques; Performance Based Education; Performance Criteria; *Predictive Ability (Testing); Predictive Measurement; *Selection; *Test Construction; *Test Validity; Training

ABSTRACT

A method of obtaining performance predictive information concerning applicants to counselor training programs, in particular, and to similar interpersonally oriented programs for other fields, in general, is presented. The rationale for and method of the development of a work sample type instrument to simulate interpersonal interaction of the counseling type are discussed, including: the development of 35 stimulus situations and the validation of the sets of responses pertaining to the situations. The results of the two methods of initial construct validation employed--Known Groups method and convergent-discriminant validity matrix--are presented. (Author)

BEST COPY AVAILABLE

ED103448

THE DEVELOPMENT OF AN INSTRUMENT TO PROVIDE PERFORMANCE PREDICTIVE INFORMATION FOR COUNSELOR TRAINEE SELECTION: PHASE I

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

by

Rory Remer, Ph.D.
Assistant Professor
Educational Foundations
University of Wyoming
January, 1973

TM004 279

BEST COPY AVAILABLE

Abstract

A method of obtaining performance predictive information concerning applicants to counselor training programs, in particular, and to similar interpersonally oriented programs for other fields, in general, is presented. The rationale for and method of the development of a work sample type instrument to simulate interpersonal interaction of the counseling type are discussed, including: the development of thirty-five stimulus situations and the validation of the sets of responses pertaining to the situations. The results of the two methods of initial construct validation employed-- "Known Groups" method and convergent-discriminant validity matrix--are presented.

THE DEVELOPMENT OF AN INSTRUMENT TO PROVIDE PERFORMANCE PREDICTIVE

INFORMATION FOR COUNSELOR TRAINEE SELECTION
January, 1973

At present, no well validated method of obtaining any clear indication of the interpersonal competence either presently possessed by or capable of being developed in an applicant to a counselor training program exists. As Wrenn (1952), Patterson (1967), and Whiteley (1968, 1969), comment an inclusion of information relevant to the potential counseling performance of an applicant is needed when making the initial selection decision. Harm can result when such information is not considered.

Patterson (1967) and Whiteley (1968, 1969), present evidence which shows that attempts to apply multiple regression techniques to the information obtained from personality and interests tests have proved, on the whole, fruitless. Equally ineffective results have been obtained while attempting to isolate characteristics which discriminate effective counselors either from those who are ineffective or from people in general.

In a review of a number of studies concerning selection procedures used for admission to counselor education programs, Remer (1972) found that traditional selection criteria such as grade point averages, GRE scores, MAT scores, years teaching experience and recommendations bear little, if any, consistent relationship to success either as a counselor in the field or even as a trainee. This information points to a need for some method of obtaining potential performance information for each applicant to be used as part of the criteria for admission to counselor training programs or other programs demanding interpersonal competence.

In order to fill the need for some indicator of interpersonal competence an attempt was made to develop an instrument which could measure such a construct. The results of the development and initial validation of the Potential Interpersonal Competence Scale (PICS) are presented. This simulated, work sample based instrument is intended to supply educators in areas demanding interpersonal competence with information presently needed but unavailable, concerning applicants to their programs so that more precise and valid selection decisions can be made.

This research may encourage others, if not to aid in the validation of this instrument, at least to attempt to develop one for the purpose of obtaining potential competence information. Competence measures are not only needed for selection decisions, but for licensing and certification as well.

Method

Cronbach (1970) states that frequently when one must assess potential performance on a complex task some sample of a subject's present performance on that task can be the best indicator of his potential--i.e., a "work sample" approach. However, as Boyd and Shimberg (1971) point out, when such a sampling can lead to injurious consequences for those involved, simulation of the task is to be preferred. Further, they indicate that the probability of obtaining a valid measurement is greatly enhanced if the behavior to be assessed is specifically identified. Certainly these statements apply to the measurement of potential interpersonal competence.

The Instrument of Development

With the comments of Cronbach (1970) and Boyd and Shimberg (1971) in mind, an instrument to measure the interpersonal competencies needed by a counselor was developed. Three aspects of the development were given

particular consideration: the formulation of the competencies desired of a counselor into specific, measurable, behavioral terms; the simulation of a counseling interview to approximate as closely and safely as possible the actual situation to be encountered in the field; and the methods by which standardized stimuli could be presented and reliable responses obtained.

In regard to the first consideration, behaviorally defined competencies outlined in the "Ability to Relate to Clients" section of the Application to the California Commission on Teacher Preparation and Licensing for Approval of Stanford University's School of Education Counseling Psychology Program (1972, p. 7) were employed in a slightly modified form. The modified competencies were:

1) Relate to the client in his own language (understanding)
2) Establish trust and confidence:
    a) Post Hoc Judgment
    b) Does not discuss confidences
3) Have rapport with client:
    a) Accurate Empathy
    b) Personal Communication
4) Look at client as an individual (do not prejudge client)
5) Build up client's confidence (in his ability to deal with his own problems)

The instrument was designed to produce a measure of each subject's interpersonal competence as defined by the above modified behaviors.

In order to simulate a counseling interview as closely as possible, actual client problem presentations were solicited from counselors.

The thirty-five problem presentations, five for each competence, were chosen from a number representative of the type encountered when working with clients ranging from high school age through adulthood.

Videotaped interviews with actors portraying clients presenting the above mentioned problems were made. These videotaped presentations provided circumstances close to the actual counseling situation without the danger inherent in actual interaction and also provided a standard stimulus to which applicants could react. A similar typescript presentation including both the problem presentation and a short non-affective, behavioral description of the client also was prepared. The differences in cost of preparation and in ease of administration between the two different modes of presentation were kept in mind, the typescript being less expensive and easier to administer. Two other modes of presentation, audiotape and timed typescript also were prepared for the purpose of comparing discriminating quality.

A multiple choice response form used with all four modes of problem presentation was prepared in accord with competencies desired of a counselor as outlined above. There were thirty-five sets of responses, five corresponding to each of the competencies as was the case with the presentations. Each set was composed of four response choices: a response designed to exemplify the competence desired to a high degree; a response which was intended to demonstrate the competence to a lesser degree; a third response designed to be diametrically opposed to the intent of the competence; and a response which was intended to be inappropriate to the counseling context but appropriate to the conversational aspects of the problem. The responses within each set then were ordered randomly to prevent any inadvertent response order biases.

The instrument, composed of typescript problem presentations and corresponding multiple choice response sets, was then administered to a panel of five behavioral counseling experts from across the country in an attempt to validate the response sets. Correctness of response was assumed to hold across mode of problem presentation. Each expert was requested to allot a total of 10 points among the four responses in each set in accord with his judgment as to the appropriateness of the responses. Each could allot the total 10 points to any alternative or any amount to each of the four alternatives as long as the total amount of points equalled 10.

Responses were designated "correct" if they met one of two criteria relative to the compilation of the points alloted to each answer by the expert panel. The criteria used were: a clear majority, 60 percent, of the points alloted, or a 50-60 percent consensus on high difficulty questions. If neither of the criteria were met, the response(s) were rewritten and resubmitted to the panel.

Of the thirty-five questions, twenty-nine response sets reached criterion on the first attempt. The remaining six qualified on the first iteration. The Hoyt method of calculating interrater reliability was employed. A value of .93 was obtained, indicating consistency in judgment of the panel.

The administration of each mode of problem presentation was designed with power considerations in mind. No attempt was made to induce a time-pressure situation. Administrators of each of the audiotape and videotape modes were instructed to allow sufficient time for all subjects to answer before presenting the next stimulus situation. The timed typescript subjects were allowed a total of 70 minutes for responding to the thirty-five

written situations. In no group did the subjects experience any difficulty with time constraints. The response time required in the audio and video groups was from 25-45 seconds per question. Most subjects in the timed typescript group completed the task within 60 minutes.

Procedures-Construct Validation

An instrument such as the one discussed here is difficult to validate since no immediate external criterion exists with which a comparison can be made. An indication is needed to determine whether the instrument is functioning well enough to warrant further investigation and/or to justify the use of longitudinal validation procedures. In order to obtain information on which to base such a decision the "Known Groups" method of construct validation (Scott 1968, p. 253) was employed.

This procedure consists of administering the instrument to a number of groups of subjects known to differ along a continuum with respect to the trait the instrument purports to measure. If the instrument does in fact distinguish among the groups in the predicted manner evidence of construct validity is obtained.

Correlations between total test score and the admissions criteria presently employed--years of teaching experience, years of counseling experience, undergraduate GPA, graduate GPA, hours of counseling course work, GRE-Quantitative score, GRE Verbal score, age, sex, and marital status--also were produced. These were used to form a partial convergent-discriminant validity matrix (Cronbach and Meehl, 1955 and Campbell and Fiske, 1959). This information added both to the knowledge of the instrument and to the knowledge of the present criteria used for admissions.

Subjects

Four groups of subjects judged to vary in their ability to relate to others were chosen. The first two groups, each composed of 20 subjects, were comprised of professional counselors with at least two years experience and practicum students with at least one semester of practicum experience. These subjects represented the high end of the ability continuum. Twenty-one students from a Foundations of Guidance course at the University of Colorado made up the middle group because they showed interest in a helping profession but were as yet untrained. This group was similar in composition to applicants to counselor training programs. Twenty-one graduate physics students constituted the low ability level group because of their self-selection into a non-person oriented profession.

The 82 subjects were randomly assigned to mode of presentation within ability level grouping. The audiovisual and the untimed typescript groups consisted of 21 subjects each, while the other two modality groups consisted of 20 subjects each. The instrument was administered, scored and the results analyzed using a two-way Analysis of Variance and the Scheffé Multiple Comparisons Procedure.

Results and Discussion

The two-way Analysis of Variance computed on the test scores produced a significant ($p < .001$) ability level effect, indicating that the instrument did indeed differentiate among the four groups of subjects. Neither a significant mode of presentation nor an ability level-by-mode of presentation interaction was evident.

Insert Table I

A series of Scheffé Multiple Comparisons were computed on the mean differences among the four groups to determine whether the differences were in the predicted directions.

Insert Table II

A total of nine comparisons were computed: Experts versus Practicum Students, Applicants versus Physics Students, Experts versus Applicants, Experts versus Physics Students, Practicum Students versus Physics Students, Practicum Students versus Applicants, Experts and Practicum Students versus Applicants, Experts and Practicum Students versus Physic Students, and Experts and Practicum Students versus Applicants and Physic Students. All comparisons proved to be significant ($p < .001$) except the comparison of Experts to Practicum Students.

Insert Table III

The graph of the relative distribution of the total tests scores for each of the four ability level groups appears in Figure 1.

Insert Figure I

Correlations between total test score and the supplementary data--years of teaching experience, undergraduate GPA, etc.--were obtained and are presented in Table IV.

Insert Table IV

A number of these results need explanation. The correlations between total test score and years of counseling experience, hours of counseling course work, GRE-Quantitative score and sex tend to be spuriously large due in part to the differences in composition between the upper ability level groups and the Physics Student Group. In particular the correlation between sex and total test score is inflated since the Physics Students were slightly over 95 percent males.

Conclusions

The development of a work sample type instrument along the lines discussed here seems not only necessary but feasible. The strong evidence of ability to discriminate among the four groups employed for validation purposes, as well as the high Hoyt reliability obtained during the testing (.84), seem to indicate that the instrument whose development has been outlined here warrants both further investigation and efforts toward longitudinal validation.

The correlation of the total test scores with the criteria presently used for admissions to counselor training programs seems to indicate that an instrument designed to measure potential interpersonal competence can add considerably to the criteria presently considered by counselor educators during the selection procedure. Hopefully, this additional information can go far in improving the procedures, making them more efficient and more valid for all those either directly or indirectly affected.

Evidence seems to indicate that the instrument whose development is discussed here can provide a much needed tool. Although results are not entirely conclusive, they are extremely encouraging. Further investigation, validation and development are needed, but a substantial advance has been made and should certainly be pursued further.

NOTE: The Potential Interpersonal Competence Scale (PICS) is copyrighted. Further information concerning the instrument itself or its development and validation, both completed and planned, can be obtained by writing the author.

References

Application to the California Commission on Teacher Preparation and Licensing for Approval of the Stanford University's School of Education Counseling Psychology Program, 1972, p. 7.

Boyd, J. L., Jr. and Shimberg, B., "Handbook of Performance Testing: A Practical Guide for Test Makers," ERIC, 1971, EDO 52220.

Campbell, D. T. and Fiske, D. W., "Convergent and Discriminant Validation by Multitrait - Multimethod Matrix," Psychological Bulletin, 1959, 56:81-105.

Cronbach, L. J., Essentials of Psychological Testing. New York: Harper and Row, 1970.

Cronbach, L. J., and Meehl, P. E., "Construct Validation in Psychological Tests," Psychological Bulletin, 1955, 52:281-302.

Patterson, C. H., "The Selection of Counselors," in Research in Counseling, J. M. Whiteley (ed.). Columbus, Ohio: Charles E. Merrill, 1967, pp. 69-102.

Remer, R., A Performance Criterion for Counselor Selection, Doctoral Dissertation: University of Colorado, 1972.

Scott, W. A., "Attitude Measurement," in The Handbook of Social Psychology, G. Lindsey and E. Aronson, (eds.), Vol. 2. Reading Massachusetts: Addison - Wesley Publishing Company, 1968, pp. 204-273.

Whiteley, J. M., "Counselor Education: A Critical Review of the Literature, 1965-1968," ERIC, 1968, EDO 26674.

Whiteley, J. M., "Counselor Education," Review of Educational Research, 1969, 2:173-187.

Wrenn, C. G., "The Selection and Education of Student Personnel Workers," Personnel and Guidance Journal, 1952, 3:9-14.

TABLE I

TWO-FACTOR 4x4 ANALYSIS OF VARIANCE
MODE OF PRESENTATION BY
ABILITY LEVEL GROUPING

Source of Variation	SS	df	MS	F
Mode	579.62	3	193.21	1.37
Level	11631.10	3	3877.03	27.55*
Mode X Level	1194.65	9	132.74	.94
Residual	9288.53	66	140.74	

* Significant $p < .001$

TABLE II

TABLE OF MEAN DIFFERENCES

Group	Expert	Practicum	Applicant	Physics
Means	115.70 (9.9)[a]	117.10 (11.5)[a]	101.05 (13.3)[a]	87.38 (12.7)[a]
Expert	0	-1.4	14.65*	28.32*
Practicum		0	16.05*	29.72*
Applicant			0	13.67*

* Significant $p < .001$

a-standard deviations in parentheses

TABLE III

SCHEFFÉ TESTS

Scheffé Comparisons

Comparison	$C_{\mu E}$	$C_{\mu PR}$	$C_{\mu AP}$	$C_{\mu PH}$	$\psi^2/\hat{\sigma}^2_{\psi}(J-1)$	F
$\psi 1$	1	-1	0	0	.14	F(1,38)
$\psi 2$	0	0	1	-1	13.94	F(1,40)*
$\psi 3$	1	0	-1	0	15.62	F(1,39)*
$\psi 4$	1	0	0	-1	64.29	F(1,39)*
$\psi 5$	0	1	0	-1	58.38	F(1,39)*
$\psi 6$	0	1	-1	0	18.75	F(1,39)*
$\psi 7$	½	½	-1	0	11.62	F(2,58)*
$\psi 8$	½	½	0	-1	41.54	F(2,58)*
$\psi 9$	½	½	-½	-½	23.88	F(3,78)*

MS_W = 140.73535

* Significant $p < .001$

TABLE IV

CORRELATIONS WITH TOTAL TEST SCORE

Variable	Correlation	Mean	S.D.
Level Value	-.67**	2.52	1.12
Quartile	.94**	2.50	1.11
Years of Teaching Experience	.13	3.18	3.49
Years of Counseling Experience	.28*	2.15	3.54
Undergraduate GPA	-.10	3.08	.40
Graduate GPA	.21	3.65	.25
Hours of Counseling Course Work	.56**	15.70	19.90
GRE-Quantitative	-.41**	595.54	134.02
GRE Verbal	.05	591.69	114.47
Age	.20	30.06	6.05
Sex (Female - 1, Male - 0)	.36**	.38	.49
Marital Status (Married = 2, Single = 1)	.14	1.59	.84

* Significant $p < .05$

** Significant $p < .001$

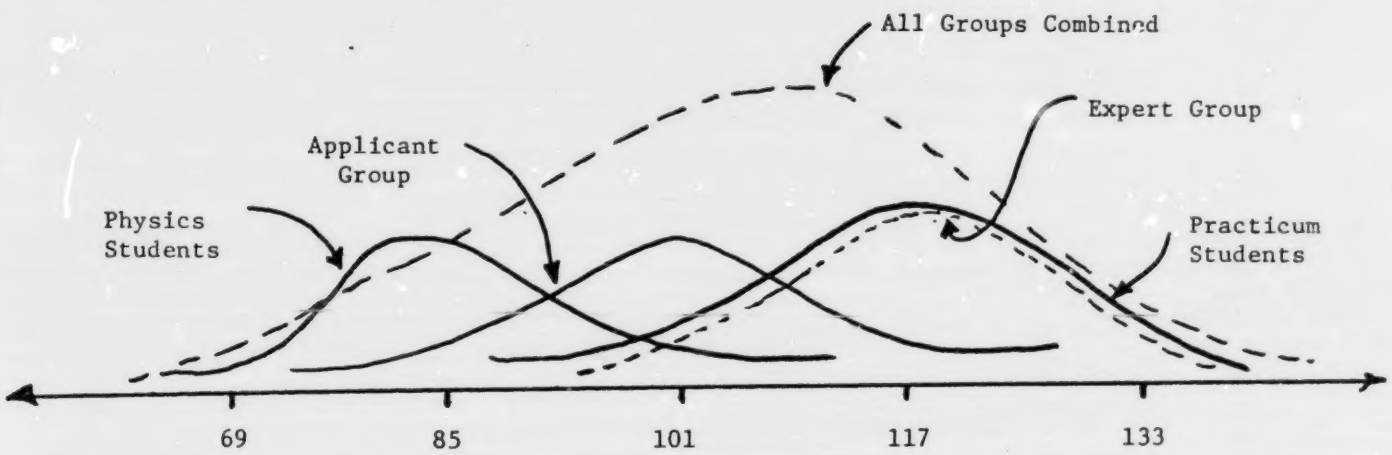

Figure I

Distribution of Test Scores by Group